Instagram, права человека и идеологическая цензура: углубленный анализ

США, ФЛОРИДА, 12 апреля 2025 г.

Введение

Instagram — одна из крупнейших в мире социальных сетей — столкнулась с растущим вниманием из-за обвинений в цензуре и дискриминации в отношении пользователей, чьи взгляды не совпадают с преобладающей политикой или идеологией платформы в отношении контента. Критики утверждают, что практика модерации контента компании иногда нарушает свободу слова и даже равносильна нарушению прав человека, особенно в случаях, связанных с консервативными, религиозными или патриотическими взглядами. В этом исследовании рассматриваются задокументированные случаи, когда Instagram блокировал или удалял контент, блокировал или подвергал теневому бану пользователей, отключал монетизацию или иным образом подавлял речь, предположительно предвзято. В нем будут рассмотрены примечательные прецеденты во время пандемии COVID-19, выборов в США 2020 года и кампании Дональда Трампа, а также конфликта между Россией и Украиной, включая случаи, когда Instagram санкционировал критиков вакцинации или мер по изоляции. Анализ также рассматривает правовые проблемы и проблемы с правами человека, ссылаясь на законы США и международные законы о свободе слова, а также исследует доказательства скрытых алгоритмов, которые могут систематически цензурировать определенный контент по политическим или иным мотивам. Целью данного отчета, составленного на основе академических и журналистских источников, является предоставление всестороннего обзора споров, связанных с модерацией контента в Instagram, и их последствий для свободы слова.

Обвинения в идеологической цензуре в Instagram

Обвинения в идеологической предвзятости модерации Instagram появились в различных сообществах. Пользователи с консервативными, религиозными или националистическими («патриотическими») убеждениями часто заявляли, что их контент несправедливо преследуется правилами платформы. Хотя официальная политика Instagram запрещает разжигание ненависти, дезинформацию и вредоносный контент, критики говорят, что эти правила применяются выборочно, чтобы заставить замолчать определенные точки зрения.

- **Консервативные голоса:**Известные политики и комментаторы правого толка уже давно утверждают, что платформы социальных сетей «незаконно подавляют консервативные взгляды» (Трамп подал в суд на Facebook, Twitter и Google, заявив о цензуре | Reuters). Например, в середине 2021 года бывший президент Дональд Трамп подал иски против Facebook/Instagram и других технологических компаний, утверждая, что они нарушили права,

гарантированные Первой поправкой, запретив его и других, — хотя эксперты по правовым вопросам отметили, что частные компании не связаны Первой поправкой (Трамп подал в суд на Facebook, Twitter и Google, заявив о цензуре | Reuters). Подавляющее большинство республиканцев (около 90%) считают, что сайты социальных сетей намеренно подвергают цензуре политические взгляды (Большинство американцев считают, что сайты социальных сетей подвергают цензуре политические взгляды). Конкретные инциденты подкрепляют эту веру: в августе 2024 года Instagram внезапно удалил несколько аккаунтов известных французских правых активистов, некоторые из которых имели десятки тысяч подписчиков, хотя они не совершили никаких правонарушений (Instagram закрывает аккаунты нескольких французских правых голосов—Европейский Консерватор). Среди запрещенных были лидеры националистических или антиправительственных групп, такие как активистка феминистского движения Элис Кордье и ассоциация поддержки осведомителей ASLA (Instagram закрывает аккаунты нескольких французских правых голосов—Европейский Консерватор). Instagram ссылался на нарушения (например, «трансфобию» в постах, критикующих трансгендерную идеологию) в качестве оправдания (Instagram закрывает аккаунты нескольких французских правых голосов—Европейский Консерватор). Пострадавшие пользователи и политики осудили чистку как «vraie purge politique» (истинную политическую чистку) и нападение на свободу слова, и даже члены французского парламента высказались (Instagram закрывает аккаунты нескольких французских правых голосов—Европейский Консерватор) (Instagram закрывает аккаунты нескольких французских правых голосов—Европейский Консерватор). Эти случаи иллюстрируют мнение о том, что модерация Instagram может непропорционально сильно влиять на консервативные или националистические взгляды.

- **Религиозные убеждения:** Религиозные консерваторы также сообщали о цензуре. В одном известном случае Instagram удалил пост баптистского служения, в котором утверждалось, что пасторами должны быть только мужчины, пометив его как разжигание ненависти (за «преследующий» контент) (Instagram удалил пост баптистского служения о том, почему пасторами должны быть только мужчины, а затем извинился | США) (Instagram удалил пост баптистского служения о том, почему пасторами должны быть только мужчины, а затем извинился | США). В посте цитировалось библейское писание (1 Тимофею 2:12–14) в поддержку традиционного взгляда на гендерные роли в служении. После того, как министерство публично обратилось с апелляцией, Instagram признал, что «ошибочно удалил этот фрагмент контента» и восстановил пост, извинившись за ошибку (Instagram удалил пост баптистского служения о том, почему пасторами должны быть только мужчины, а затем извинился | США). Инцидент (который министерство охарактеризовало как «цензурирование библейской истины») предполагает, что автоматизированные фильтры или чрезмерно ревностные модераторы могут ошибочно запретить контент, выражающий основные религиозные позиции по социальным вопросам. Аналогичным образом в 2020 году Instagram объявил, что запретит контент, пропагандирующий «конверсионную терапию» (практику, против которой выступают защитники ЛГБТК) как форму языка ненависти (Facebook и Instagram запрещают контент, пропагандирующий «конверсионную терапию»). Хотя такая политика направлена на защиту пользователей, она привела к жалобам некоторых религиозных групп на то, что выражение традиционных убеждений относительно сексуальности может привести к несправедливой цензуре.
- **Патриотическое и политическое содержание:** Самопровозглашенные «патриоты» и популистские группы подняли тревогу по поводу теневых банов или подавления подписчиков в Instagram. В США некоторые консервативные влиятельные лица утверждают, что Instagram тайно ограничивает охват проамериканского или националистического контента, особенно если он бросает вызов преобладающим политическим нарративам. Например, во время избирательного сезона 2020 года пользователи утверждали, что посты в поддержку тогдашнего президента Трампа или подвергающие сомнению результаты выборов были понижены или скрыты. (Facebook, материнская компания Instagram, действительно приняла агрессивные меры по пресечению дезинформации о выборах, которые в основном были нацелены на нарративы, распространенные в правых кругах (Трамп подал в суд на Facebook, Twitter и Google, заявив о цензуре | Reuters).) На международном уровне картина повторяется: массовые блокировки французских правых аккаунтов в августе 2024 года были оправданы

Instagram в соответствии с его политикой, но критики отметили, что ни один из запрещенных материалов не был незаконным во Франции (Instagram закрывает аккаунты нескольких французских правых голосов—Европейский Консерватор). Пострадавшие пользователи, многие из которых идентифицируют себя как патриотически настроенные защитники французских традиций, утверждали, что Instagram заглушает голоса, «которые защищают Францию и правду» (Instagram закрывает аккаунты нескольких французских правых голосов—Европейский Консерватор). Даже юмор не был пощажен: французская страница пародийного мема «Une Bonne Droite» (дословно «хороший правый [хук]», каламбур на тему правого крыла) была удалена, несмотря на то, что была сатирической (Instagram закрывает аккаунты нескольких французских правых голосов—Европейский Консерватор). Эти примеры подчеркивают, почему многие идеологически правые или патриотически настроенные пользователи чувствуют себя систематически дискриминированными на платформе.

Следует отметить, что вопрос предвзятости является предметом горячих споров. Исследования и заявления платформы иногда противоречат этим утверждениям. Руководители Facebook/Instagram отрицают преднамеренную предвзятость (генеральный директор Instagram Адам Моссери заявил в 2020 году, что «теневой бан — это не вещь» с точки зрения таргетирования определенных аккаунтов (Теневой бан реален: как социальные сети решают, кому заткнуть рот - The Washington Post)), и некоторые внешние исследования не находят никаких доказательств системной антиконсервативной предвзятости (Утверждения об антиконсервативной предвзятости со стороны компаний социальных сетей беспочвенны...). Действительно, данные Facebook часто показывают, что консервативные эксперты генерируют самую высокую вовлеченность. Тем не менее, восприятие имеет значение — и опрос Pew Research показал, что подавляющее большинство пользователей-республиканцев считают, что цензура имеет место (Большинство американцев считают, что сайты социальных сетей подвергают цензуре политические взгляды). В следующих разделах описаны конкретные случаи удаления контента и санкций в отношении аккаунтов, которые способствовали такому восприятию.

Практики модерации контента: от банов до теневых банов

Instagram использует ряд инструментов модерации, которые могут повлиять на видимость пользователей и их способность делиться контентом. К ним относятся прямые карательные действия (например, удаление постов или блокировка аккаунтов), а также алгоритмические методы (сокращение охвата поста или «теневой бан»). Понимание этих механизмов имеет решающее значение для оценки заявлений о цензуре:

- **Удаление контента и запреты:**Наиболее очевидной формой модерации является удаление отдельных постов или постоянная блокировка аккаунтов, которые нарушают Правила сообщества Instagram. Например, Instagram удалит посты, содержащие наготу, сцены насилия, откровенные высказывания, разжигающие ненависть, или опасную дезинформацию. В громких случаях Instagram блокировал влиятельные аккаунты за распространение ложной информации — например, в феврале 2021 года Instagram заблокировал аккаунт Роберта Ф. Кеннеди-младшего после неоднократных нарушений его политики дезинформации о COVID-19 (Instagram заблокировал РФК-младшего за распространение дезинформации о вакцинах : Обновления о коронавирусе : NPR). Представитель Facebook объяснил, что «мы удалили этот аккаунт за неоднократный обмен опровергнутыми утверждениями о коронавирусе или вакцинах». (Instagram заблокировал РФК-младшего за распространение дезинформации о вакцинах : Обновления о коронавирусе : NPR). РФК-младший, известный скептик вакцинации, таким образом потерял свою платформу для того, что он считал «альтернативными» взглядами на здоровье, иллюстрируя, как контент, который считается вредным по общепринятым стандартам, полностью удаляется. Аналогичным образом Instagram (как часть Meta) заблокировал аккаунт тогдашнего президента Дональда Трампа в

январе 2021 года за посты, предположительно подстрекающие к насилию (бунт в Капитолии) (Трамп подал в суд на Facebook, Twitter и Google, заявив о цензуре | Reuters) — один из самых значимых запретов, вызвавший бурные дебаты о полномочиях частных платформ.

- **Демонетизация и ограничения функций:**Помимо удаления, Instagram может наложить санкции на пользователей, отключив определенные функции. Аккаунты, которые неоднократно публикуют контент, нарушающий правила, могут потерять доступ к монетизации (например, им может быть запрещено показывать рекламу, партнерские ссылки или другие функции получения дохода). Им также может быть запрещено появляться в поисковых подсказках или на страницах Explore, что серьезно сократит их аудиторию. Эти меры часто менее заметны публично, но создатели замечают их влияние. Например, многочисленные инфлюенсеры сообщали о внезапном падении охвата после того, как их пометили за дезинформацию или нарушение политики, что фактически привело к потере дохода от рекламы. (Meta обычно не публикует информацию о том, кого она демонетизирует, но в ее политике указано, что аккаунты, нарушающие правила, могут быть ограничены в рекламе или других инструментах получения дохода ().)
- **«Теневой бан» (алгоритмическое подавление):**Теневой бан относится к скрытому снижению видимости пользователя — его посты или комментарии показываются гораздо меньшему количеству людей или полностью скрываются без какого-либо уведомления пользователя. Instagram долго отрицал использование настоящих «теневых банов», но доказательства указывают на то, что платформа действительно использует подобную тактику. В одном расследовании The Markup создал тестовые аккаунты и обнаружил, что когда определенные комментарии были помечены (как спам или домогательства), Instagram делал их невидимыми для всех других пользователей, продолжая при этом показывать их автору — классический симптом теневого бана (Как мы исследовали теневой бан в Instagram – The Markup) (Как мы исследовали теневой бан в Instagram – The Markup). В этих случаях пользователь не получает уведомления о том, что его комментарий заблокирован. Instagram признал, что некоторый контент не будет рекомендован или будет понижен в рейтинге, если он будет признан пограничным или если у аккаунта есть история нарушения правил (иногда это называется «уменьшенное распространение»). Пользователи часто предполагают, что они были подвергнуты теневому бану, когда их активность резко падает без какой-либо очевидной причины. The Washington Post описывает теневой бан как «форму онлайн-цензуры, при которой вам все еще разрешено говорить, но вряд ли кто-то вас слышит», реализуемую с помощью «невидимых алгоритмов» без прозрачности (Теневой бан реален: как социальные сети решают, кому заткнуть рот - The Washington Post) (Теневой бан реален: как социальные сети решают, кому заткнуть рот - The Washington Post). Эта непрозрачная практика затронула различные группы: консервативные активисты утверждают, что их посты на такие темы, как фальсификация выборов, были тихо скрыты, в то время как некоторые прогрессивные активисты обнаружили, что посты о расовой справедливости внезапно стали пользоваться меньшей популярностью (Теневой бан реален: как социальные сети решают, кому заткнуть рот - The Washington Post) (Теневой бан реален: как социальные сети решают, кому заткнуть рот - The Washington Post). Отсутствие ясности — в содержании или в алгоритме? — оставляет многих «жертв, полных неуверенности в себе, граничащей с паранойей» (Теневой бан реален: как социальные сети решают, кому заткнуть рот - The Washington Post).
- **Манипулирование хэштегами и подписчиками:**Instagram также может подавлять определенные хэштеги или поисковые запросы, а периодические чистки фейковых аккаунтов могут привести к потере подписчиков пользователями. Например, в январе 2025 года пользователи заметили, что поиск политически мотивированных хэштегов, таких как «#Democrat», «#Republicans» или «#Jan6th», возвращал сообщение о том, что «мы скрыли эти результаты» (Пользователи Instagram сталкиваются с заблокированными политическими хэштегами в результатах поиска) (Пользователи Instagram сталкиваются с заблокированными политическими хэштегами в результатах поиска). Это вызвало обвинения в политической предвзятости на Reddit и в других местах, хотя представитель Instagram заявил, что это технический сбой, затронувший многие термины (и пообещал исправить) (Пользователи Instagram сталкиваются с заблокированными политическими хэштегами в результатах поиска). Что касается подписчиков, некоторые откровенные аккаунты утверждают, что

> Instagram принудительно отписывает их фанатов, чтобы ограничить их охват. Хотя Instagram этого не подтвердил, внезапные большие падения числа подписчиков часто объясняются тем, что Instagram отсеивает ботов или неактивные аккаунты из системы (Кто-нибудь еще только что потерял кучу подписчиков в сфере ... - Reddit). Однако эти рутинные чистки могут быть неверно истолкованы пользователями как целенаправленные действия. Чистый эффект этих более тонких инструментов — блокировка хэштегов, алгоритмическая отписка, понижение рейтинга — является формой мягкой цензуры: контент не запрещается напрямую, но его становится намного сложнее найти или распространить.

(Instagram удалил пост баптистского служения о том, почему пасторами должны быть только мужчины, а затем извинился | США) Рисунок: Логотип Instagram. Критики утверждают, что политика платформы в отношении контента иногда применяется таким образом, что непропорционально подавляет определенные точки зрения. (Instagram удалил пост баптистского служения о том, почему пасторами должны быть только мужчины, а затем извинился | США) (Теневой бан реален: как социальные сети решают, кому заткнуть рот - The Washington Post)

В целом, модерация Instagram варьируется от жестких удалений (удаление постов/аккаунтов) до мягкого подавления (сокрытие контента в алгоритмических тенях). Для пользователей, которые сталкиваются с этими действиями — особенно без четкого объяснения — это действительно ощущается как цензура. Следующие примеры иллюстрируют, как эти меры применялись во время недавних громких событий и споров, часто разжигая дебаты о свободе слова и силе платформы.

Пример 1: Пандемия COVID-19 — дезинформация о здоровье против свободы слова

Во время глобальной пандемии COVID-19 (2020–2021) Instagram (как и другие платформы) резко ужесточил свою политику по борьбе с дезинформацией о вирусе, вакцинах и мерах общественного здравоохранения. Хотя это было обусловлено императивами общественного здравоохранения, это привело к многочисленным столкновениям с людьми, которые публиковали противоречивые или скептические взгляды на вакцины, карантины или методы лечения. Эти пользователи часто рассматривали вмешательства Instagram как цензуру законного инакомыслия.

Меры по борьбе с антипрививочным контентом: Instagram явно запретил множество ложных или вводящих в заблуждение заявлений о COVID-19 и вакцинах на том основании, что такая дезинформация может нанести вред общественному здравоохранению. На практике это означало удаление тысяч постов и блокировку аккаунтов, включая аккаунты известных деятелей антипрививочного движения. Удаление аккаунта Роберта Ф. Кеннеди-младшего в феврале 2021 года стало знаковым примером. РФК-младший, давний критик вакцин, собрал более 800 000 подписчиков, публикуя контент, сеющий сомнения в безопасности вакцин и продвигающий теории заговора (например, что Билл Гейтс использовал вакцины для получения прибыли или контроля) (Instagram заблокировал РФК-младшего за распространение дезинформации о вакцинах : Обновления о коронавирусе : NPR) (Instagram заблокировал РФК-младшего за распространение дезинформации о вакцинах : Обновления о коронавирусе : NPR). Instagram в конце концов забанил его за неоднократные нарушения, заявив, что он неоднократно делился опровергнутыми утверждениями, которые подрывали доверие к вакцинам от COVID-19 (Instagram заблокировал РФК-младшего за распространение дезинформации о вакцинах : Обновления о коронавирусе : NPR). Аналогичным образом Instagram закрыл популярные антипрививочные хэштеги (такие как #vaccineskill или #vaccinescauseautism), чтобы остановить распространение таких постов (Как антипрививочники обходят новый контроль хэштегов в Instagram). В то время как эксперты в области общественного здравоохранения приветствовали эти шаги, оппоненты утверждали, что технологические компании фактически подавляют точку зрения меньшинства, отмечая, что некоторый контент (например, истории о побочных эффектах вакцин) был помечен как

«дезинформация», несмотря на то, что содержал реальные личные истории, поскольку противоречил официальной информации.

Цензура или безопасность?Защитники Instagram утверждают, что удаление явно ложной информации (например, утверждений о том, что вакцины от COVID имплантируют микрочипы) — это не цензура, а ответственная модерация. Однако возникла серая зона вокруг «критиков вакцинации или санитарных ограничений» — людей, которые, например, сомневались в эффективности карантинов, масочного режима или вакцинации, не обязательно распространяя ложь. Многие такие пользователи обнаружили, что их посты также были помечены или удалены. Некоторые сообщили, что простое выражение скептицизма по поводу скорости одобрения вакцин или пропаганда альтернативных методов лечения приводили к удалению. Это создало пугающий эффект, при котором даже добросовестные дебаты были подавлены из-за страха модерации. Проблема достигла апогея, когда выяснилось, что чиновники правительства США оказали давление на платформы социальных сетей, чтобы те ужесточили меры в отношении постов о COVID-19. В августе 2024 года генеральный директор Meta Марк Цукерберг сообщил в письме, что «высшие должностные лица администрации Байдена… неоднократно оказывали давление на наши команды в течение месяцев, чтобы они подвергали цензуре определенный контент о COVID-19, включая юмор и сатиру» (Цукерберг заявил, что администрация Байдена оказала давление на Meta, чтобы тот подверг цензуре контент о COVID-19 | Reuters) (Цукерберг заявил, что администрация Байдена оказала давление на Meta, чтобы тот подверг цензуре контент о COVID-19 | Reuters). Сюда входят жалобы Белого дома в 2021 году на то, что Facebook/Instagram «убивают людей», не удаляя достаточно дезинформации о вакцинах (Цукерберг заявил, что администрация Байдена оказала давление на Meta, чтобы тот подверг цензуре контент о COVID-19 | Reuters). Признание Цукерберга, сделанное Конгрессу, подразумевало, что некоторые решения по модерации контента во время пандемии были обусловлены не только частной политикой, но и влиянием правительства — динамичные критики называют это «отмыванием цензуры» (когда правительства косвенно цензурируют через частные компании, чтобы избежать проверки Первой поправки). Действительно, в 2023 году в федеральном иске (Миссури против Байдена) утверждалось, что американские чиновники вступили в сговор с технологическими фирмами с целью подавления антилокдауновых или антивакцинальных высказываний, что подняло новые вопросы о Первой поправке. Пока это дело продолжается, разоблачение подчеркивает, как антиистеблишментские взгляды на COVID-19 агрессивно контролировались в Instagram, возможно, в ущерб открытому дискурсу (Цукерберг заявил, что администрация Байдена оказала давление на Meta, чтобы тот подверг цензуре контент о COVID-19 | Reuters) (Цукерберг заявил, что администрация Байдена оказала давление на Meta, чтобы тот подверг цензуре контент о COVID-19 | Reuters).

Результирующая реакция:Пандемия подчеркнула напряженность между общественной безопасностью и свободой слова. Instagram, вероятно, предотвратил распространение некоторой опасной дезинформации, но его тактика также спровоцировала движение, осуждающее «цензуру крупных технологических компаний». Инфлюенсеры, которые были деплатформированы в Instagram, перешли на альтернативные платформы (например, Telegram или Parler) с криками о преследовании. Даже некоторые ученые и журналисты стали обеспокоены — не потому, что они одобряли дезинформацию, а потому, что законные вопросы или мнения меньшинства легко попадали в сеть Instagram. Оглядываясь назад, можно признать ошибки в правоприменении: совет по надзору Meta рассмотрел несколько решений об удалении, связанных с COVID, иногда отменяя их как необоснованные. Например, совет по надзору восстановил пост, удаленный за цитирование спорных комментариев главы государства о COVID, отметив, что обсуждение политики правительства — даже если оно спорное — является защищенным политическим выражением в соответствии с международными стандартами прав человека (). Это свидетельствует о том, что модерация Instagram в разгар пандемии иногда перегибала палку. С тех пор этот опыт привел к тому, что Meta стала принимать больше прозрачности (например, пропускать некоторые решения через проверяющих факты и предлагать метки вместо удаления пограничного контента). Тем не менее, эпоха COVID-19 является случаем, когда борьба Instagram с дезинформацией напрямую столкнулась с правами людей высказывать альтернативные точки зрения, что подпитывает продолжающееся недоверие среди этих сообществ.

Пример 2: Выборы в США и политическая речь – эпоха Трампа

Социальные сети сыграли огромную роль в политической нестабильности конца 2010-х и начала 2020-х годов. Instagram, как часть экосистемы Meta, был вовлечен в усилия по контролю за лживыми фактами, связанными с выборами, и экстремистским контентом, который сторонники рассматривали как защиту демократии, а критики считали политически мотивированной цензурой. Сага вокруг контента Дональда Трампа и связанных с ним движений (таких как Stop the Steal и QAnon) иллюстрирует этот конфликт.

Дезинформация о выборах и ее модерирование:В преддверии президентских выборов в США 2020 года Instagram ввел более строгие правила против ложных заявлений о процедурах голосования, результатах выборов и делегитимации результатов. Контент, который считался распространяющим необоснованные обвинения в широкомасштабном мошенничестве на выборах — повествование, активно продвигаемое Трампом и его сторонниками, — был помечен или удален. Например, хэштеги, связанные с теориями заговора на выборах (например, #StopTheSteal), были заблокированы в Instagram и Facebook за нарушение правил гражданской честности. Такие вмешательства усилились после выборов: когда Трамп и его союзники продолжили утверждать, что выборы были «украдены», Instagram удалил некоторые из этих постов и предотвратил появление определенных фраз в трендах. 6 января 2021 года, когда в Капитолии США вспыхнули беспорядки, Meta предприняла беспрецедентный шаг, приостановив на неопределенный срок действия аккаунтов президента Трампа в Facebook и Instagram, сославшись на риск дальнейшего насилия (Трамп подал в суд на Facebook, Twitter и Google, заявив о цензуре | Reuters). Этот запрет заставил действующего президента США замолчать на двух крупнейших мировых площадках и был широко встречен противниками Трампа, а его сторонники осудили его как вопиющую политическую цензуру.

Утверждения об антиконсервативной предвзятости:Инцидент с Трампом стал поводом для обсуждения для тех, кто утверждает, что элита Кремниевой долины предвзято относится к консерваторам. Сам Трамп неоднократно заявлял, что «правые голоса подвергаются цензуре» на таких платформах, как Facebook и Instagram (Пользователи Instagram сталкиваются с заблокированными политическими хэштегами в результатах поиска) (несмотря на то, что данные показывают, что консервативный контент часто доминирует над вовлеченностью (Пользователи Instagram сталкиваются с заблокированными политическими хэштегами в результатах поиска)). В июле 2021 года Трамп и группа других лиц подали коллективный иск против Facebook (материнской компании Instagram), Twitter и YouTube, обвинив их в цензурировании консерваторов и потребовав восстановления их аккаунтов (Трамп подал в суд на Facebook, Twitter и Google, заявив о цензуре | Reuters) (Трамп подал в суд на Facebook, Twitter и Google, заявив о цензуре | Reuters). В иске даже требовалось признать недействительным раздел 230 Закона о благопристойности в коммуникациях (Трамп подал в суд на Facebook, Twitter и Google, заявив о цензуре | Reuters) — закон США, который предоставляет платформам иммунитет за добросовестное модерирование контента. Юристы отметили иронию: раздел 230 фактически защищает возможность платформ удалять контент (то есть его закрытие может привести к тому, что компании будут реже размещать спорные высказывания, а не чаще). В конечном итоге судья отклонил дело Трампа, подтвердив, что частные компании не связаны требованиями Первой поправки о свободе слова и что никакие доказательства не превращают их в «государственных субъектов» в этом контексте (Трамп подал в суд на Facebook, Twitter и Google, заявив о цензуре | Reuters). Однако политическое давление со стороны правых сохранялось, что привело к тому, что такие штаты, как Флорида и Техас, приняли законы, запрещающие цензуру в социальных сетях политических кандидатов или точек зрения. (Закон Флориды был заблокирован в суде как неконституционный, а закон Техаса находится на рассмотрении в суде (Трамп подал в суд на Facebook, Twitter и Google, заявив о цензуре | Reuters).)

Восстановление Трампа в должности и юридическая победа:Удивительным образом к 2023–2024 годам Meta начала пересматривать некоторые запреты. Согласно внутренней политике (и руководству независимого Совета по надзору), они объявили, что отстранение Трампа будет снято через два года (т. е. в 2023 году) с «ограничениями» на его будущие посты. Затем в январе 2025 года Meta согласилась урегулировать иск Трампа, прежде чем он зайдет дальше, — заплатив Трампу 25 миллионов долларов, чтобы разрешить его претензии о том, что Facebook/Instagram занимались «недопустимой цензурой», забанив его (Meta согласилась выплатить Трампу 25 миллионов долларов для урегулирования иска из-за блокировок Facebook и Instagram: NPR). Это урегулирование, по которому, как сообщается, крупная сумма была направлена в Президентскую библиотеку Трампа, было воспринято как «значительная победа Трампа» и вызвало удивление во всем политическом спектре (Meta согласилась выплатить Трампу 25 миллионов долларов для урегулирования иска из-за блокировок Facebook и Instagram: NPR). Некоторые критики (например, сенатор Элизабет Уоррен) раскритиковали выплату как «взятку… сигнал о том, что коррупция — это название игры», предполагая, что Мета может ожидать взамен политических услуг (Meta согласилась выплатить Трампу 25 миллионов долларов для урегулирования иска из-за блокировок Facebook и Instagram: NPR). Независимо от мотивов, соглашение фактически положило конец спору с Трампом без судебного прецедента, но оно признало политическую реальность, что Meta необходимо было наладить отношения с правыми. К концу 2024 года Марк Цукерберг даже приказал смягчить некоторые правила модерации — например, отказаться от строгой проверки фактов в пользу «заметок сообщества» и снова разрешить рекомендовать политический контент — в том, что многие увидели как попытку противостоять обвинениям в антиконсервативной предвзятости по мере приближения выборов 2024 года (Пользователи Instagram сталкиваются с заблокированными политическими хэштегами в результатах поиска) (Пользователи Instagram сталкиваются с заблокированными политическими хэштегами в результатах поиска). Результатом стал едва заметный сдвиг: Instagram сейчас сравнительно менее навязчив в плане политических высказываний, чем в 2020 году, что иллюстрирует, как общественное и юридическое противодействие может влиять на политику платформы.

Подводя итог, можно сказать, что обращение Instagram с политическим контентом эпохи Трампа демонстрирует сложное взаимодействие принципа, давления и прибыли. Изначально платформа заняла жесткую позицию в отношении того, что она считала опасной ложью (кульминацией чего стал запрет Трампа), совпадая с мнением многих пользователей о необходимости модерации. Однако негативная реакция консерваторов и угроза регулирующих/юридических действий заставили пересмотреть свое мнение. Этот эпизод подчеркивает, почему дебаты по поводу модерации в социальных сетях часто поляризуются по партийным линиям: одна сторона требует более жестких мер против «дезинформации», а другая называет это «цензурой» их мировоззрения. Instagram оказался прямо в центре этого национального спора.

Пример 3: Геополитические конфликты и цензура – Украина и не только

Проблемы модерации Instagram не ограничиваются политикой США или дебатами о здравоохранении; они также распространяются на международные конфликты, где политика платформы иногда, по-видимому, благоприятствует повествованию одной из сторон, вызывая опасения цензуры в глобальном масштабе. Два важных контекста — война между Россией и Украиной (с 2022 года по настоящее время) и конфликт между Израилем и Палестиной (2023 года), оба из которых привели к необычным решениям Instagram по модерации контента, которые вызвали критику.

Война на Украине – предвзятость в политике:В марте 2022 года, вскоре после вторжения России на Украину, внутренние метасообщения (позже подтвержденные Reuters) выявили шокирующее временное изменение политики: Facebook и Instagram решили разрешить пользователям в некоторых

странах публиковать призывы к насилию в отношении российских солдат и лидеров в контексте войны (Facebook разрешает публиковать военные посты, призывающие к насилию против российских оккупантов | Reuters) (Facebook разрешает публиковать военные посты, призывающие к насилию против российских оккупантов | Reuters). Обычно заявления типа «смерть [людям]» нарушают правила о разжигании ненависти или подстрекательстве. Но Meta сделала исключение, разрешив такие заявления, как «смерть российским оккупантам» или даже призывы к смерти президента Путина, если они не являются реальными угрозами (Facebook разрешает публиковать военные посты, призывающие к насилию против российских оккупантов | Reuters). Это было оформлено как разрешение украинцам и их соседям выражать свой гнев агрессору – по сути, занимая идеологическую позицию, которая поддерживала оборону Украины. Последствия были немедленными: правительство России раскритиковало Мету за «экстремистскую деятельность» (Facebook разрешает публиковать военные посты, призывающие к насилию против российских оккупантов | Reuters) и в течение нескольких дней полностью запретили Instagram в России, отключив десятки миллионов пользователей (Россия запретила Facebook и Instagram из-за закона об «экстремизме») (Россия ограничивает социальные сети, пока продолжается война в Украине - NPR). С одной стороны, разрешение Instagram было проявлением солидарности с нацией, подвергшейся нападению; с другой стороны, это продемонстрировало двойные стандарты в модерации контента — язык вражды допускался, когда был направлен на политически неблагоприятную группу (российские военные/руководство), что было бы немыслимо в других контекстах. И наоборот, любой контент с российской стороны, который оправдывал войну или рассматривался как пропаганда, агрессивно понижался в рейтинге или удалялся Meta, особенно после того, как западные правительства и ЕС призвали к ужесточению мер в отношении российских государственных СМИ. Instagram пометил аккаунты российских СМИ, связанных с государством, и, в соответствии с санкциями ЕС, заблокировал такие издания, как RT и Sputnik, в странах ЕС (форма принудительной цензуры этих источников).

Скрытая цензура в контенте, связанном с конфликтом:Помимо явных изменений политики, алгоритмы Instagram тихо боролись с контентом, связанным с войной. Расследование The Markup, проведенное в 2024 году, показало, что Instagram «сильно понижал рейтинг» даже неграфических изображений войны, удалял некоторые подписи к публикациям без предупреждения и скрывал многочисленные комментарии — часто без уведомления пользователей (Пониженный, удаленный и отклоненный: в Instagram есть не только теневой бан – The Markup). Он также подавлял хэштеги непоследовательными способами. Примечательно, что пользователи, которые пытались публиковать посты о войне (например, делились изображениями или репортажами из Украины), иногда обнаруживали, что их посты необъяснимым образом получали гораздо меньше просмотров, что означало, что алгоритм пометил их. В некоторых случаях комментарии о войне удалялись и помечались как «спам» (тем самым лишая пользователя возможности подать апелляцию, поскольку Instagram не допускает апелляций на удаления, классифицированные как спам) (Пониженный, удаленный и отклоненный: в Instagram есть не только теневой бан – The Markup). Это повлияло на комментарии с обеих сторон, но активисты жаловались, что важные документы — такие как сообщения украинских мирных жителей о зверствах — могли быть похоронены автоматическими фильтрами, не настроенными на военный сценарий (Создано для мирного времени, а не для войны: как Украина заставляет компании...). Системы модерации Instagram, как утверждается, были «разработаны для мирного времени, а не для войны», что привело к чрезмерной фильтрации конфликтного контента (Создано для мирного времени, а не для войны: как Украина заставляет компании...). Компании пришлось скорректировать некоторые алгоритмы после критики украинского правительства, что платформы удаляют доказательства российских военных преступлений по ошибочным правилам контента (например, политикам в отношении графического насилия). Этот сценарий подчеркивает, что цензура может возникать не только из-за преднамеренной предвзятости, но и из-за плохо откалиброванных алгоритмов, которые ошибаются в сторону удаления во время кризисов.

Израильско-палестинский конфликт – одностороннее замалчивание:Другой яркий пример произошел во время войны Израиля и ХАМАС в октябре 2023 года. Многочисленные сообщения указывают на то, что Instagram и Facebook занимались систематическим удалением контента,

поддерживающего палестинцев или критикующего военные действия Израиля. Human Rights Watch расследовала и задокументировала более 1000 случаев удаления или подавления контента в период с октября по ноябрь 2023 года, «из которых 1049 касались мирного контента в поддержку Палестины… и только 1 касался контента в поддержку Израиля». (Нарушенные обещания Меты: системная цензура палестинского контента в Instagram и Facebook | HRW). Сюда входили сообщения о нарушениях прав человека, протестах или сочувствии гражданам Газы, которые удалялись или скрывались в крайне несоразмерных количествах (Нарушенные обещания Меты: системная цензура палестинского контента в Instagram и Facebook | HRW) (Нарушенные обещания Меты: системная цензура палестинского контента в Instagram и Facebook | HRW). HRW назвала это «волной усиления цензуры», которая «заставила замолчать голоса в поддержку Палестины» во всем мире (Нарушенные обещания Меты: системная цензура палестинского контента в Instagram и Facebook | HRW). Одной из причин, вероятно, было агрессивное применение правил против восхваления террористических организаций — ХАМАС забанен на Meta — но сетевое освещение модерации ИИ было слишком широким, улавливая безобидные речи самих палестинцев о своих страданиях. Instagram также «беспорядочно подавлял» хэштеги, такие как #FreePalestine или #CeasefireForGaza, согласно сообщениям пользователей, и часто помечал посты как деликатные или понижал их рейтинг, делая их менее заметными (Пониженный, удаленный и отклоненный: в Instagram есть не только теневой бан – The Markup). Общественный резонанс от журналистов, активистов и даже некоторых сотрудников Meta побудил Надзорный совет объявить дело по этому вопросу. Meta ответил, списав некоторые проблемы на ошибки (например, сбой, из-за которого истории о войне временно отображались ниже в лентах) и пообещав исправить ошибки. Тем не менее, сохраняется восприятие того, что Instagram занял чью-то сторону — намеренно или нет — подавляя повествование одной из сторон в международном конфликте. Это не только подрывает нейтралитет платформы, но и, как отметила HRW, нарушает обязанность уважать свободу выражения мнений в соответствии с международными стандартами прав человека (Нарушенные обещания Меты: системная цензура палестинского контента в Instagram и Facebook | HRW).

В этих геополитических случаях модерация Instagram проверялась на быстро меняющемся, эмоционально заряженном контенте. Действия платформы, будь то действия, обусловленные политикой (Украина) или ошибками (Палестина), в конечном итоге вызвали обвинения в цензуре и предвзятости. Они показывают, что «официальная идеология» Instagram (которая соответствует западным либеральным взглядам и соответствует западным государственным нормам) может вступать в конфликт с альтернативными точками зрения — будь то российский патриот, защищающий действия своей страны, или голос палестинца, подчеркивающий их бедственное положение. Когда эти взгляды подавляются, это подпитывает аргументы о том, что Instagram нарушает дух свободного выражения за пределами границ США, даже если юридически Первая поправка не распространяется на частную компанию.

Правовые аспекты и вопросы прав человека

Напряжение между методами модерации Instagram и правами пользователей на свободу слова побудило обратиться к правовым рамкам, как американским, так и международным, которые защищают свободу слова. Однако в основе лежит парадокс: Instagram является частной корпорацией и, таким образом, не связан напрямую Первой поправкой к Конституции США (которая ограничивает только государственную цензуру) (Свобода слова в социальных сетях: полное руководство - Freedom Forum). Фактически, как подтвердили суды США, сами компании социальных сетей имеют право, закрепленное в Первой поправке, решать, какие высказывания разрешать на своих платформах — по сути, это их собственная свобода выражения мнений как частных лиц (Трамп подал в суд на Facebook, Twitter и Google, заявив о цензуре | Reuters). Это означает, что пользователи не могут успешно подать в суд на Instagram за нарушение их конституционных прав на свободу слова, как показал отклоненный иск Трампа.

Тем не менее, ведутся юридические и нормативные баталии за определение границ власти платформы:

- **Закон США:**Хотя Первая поправка не применяется к решениям Instagram о контенте, в игру вступают другие законы и доктрины. Раздел 230 Закона о благопристойности в коммуникациях (47 USC § 230) особенно важен. Раздел 230 предусматривает, что онлайн-платформы не несут ответственности за контент, размещенный пользователями, и, что важно, они также не несут ответственности за «добросовестное» удаление контента. Этот закон, называемый «основой интернета», позволил Instagram удалять или ограничивать контент, который он считает нежелательным, не опасаясь судебного преследования за предвзятость (Трамп подал в суд на Facebook, Twitter и Google, заявив о цензуре | Reuters). Критики раздела 230 (включая некоторых консерваторов) утверждают, что он дает технологическим компаниям слишком много полномочий для цензуры, и они лоббируют его изменение или отмену (Трамп подал в суд на Facebook, Twitter и Google, заявив о цензуре | Reuters). До сих пор Раздел 230 остается в силе, и попытки использовать законные средства, чтобы заставить Instagram размещать определенные высказывания (например, закон Флориды о социальных сетях, предусматривающий наказание за удаление политических кандидатов из платформы), столкнулись с конституционными проблемами, поскольку суды постановили, что правительство не может заставить частную платформу размещать высказывания (Трамп подал в суд на Facebook, Twitter и Google, заявив о цензуре | Reuters). Таким образом, согласно действующему законодательству США, Instagram имеет широкие возможности модерировать контент по своему усмотрению, но именно эту свободу и оспаривают некоторые законодательные органы штатов и политики.
- **Международное право прав человека:**Свобода слова также защищена международными инструментами, которые устанавливают глобальные нормы. Статья 19 Всеобщей декларации прав человека (1948) провозглашает право «искать, получать и распространять информацию и идеи любыми средствами и независимо от государственных границ». Аналогичным образом, статья 19 Международного пакта о гражданских и политических правах (МПГПП) (обязательный договор, ратифицированный США) гарантирует каждому право на свободу слова, за исключением узких ограничений (таких как подстрекательство к насилию или соображения национальной безопасности) (Что такое свобода выражения мнения? - СТАТЬЯ 19). Однако эти положения применяются к правительствам — они обязывают государства не осуществлять незаконную цензуру. Instagram, как субъект частного сектора, не является напрямую носителем обязательств по этим договорам. Однако растет ожидание того, что компании будут добровольно соблюдать стандарты прав человека. Руководящие принципы ООН в сфере бизнеса и прав человека призывают корпорации соблюдать права человека в своей деятельности, и Meta действительно приняла корпоративную политику в области прав человека, обязуясь соблюдать такие принципы, как свобода выражения мнений (). На практике Совет по надзору Meta явно использует международное право в области прав человека в качестве ориентира при рассмотрении сложных дел о модерации контента. Например, при оценке того, было ли удаление надлежащим, Совет спрашивает, было ли решение «необходимым и соразмерным» в соответствии со стандартами статьи 19 МПГПП () (). В некоторых случаях Совет находил действия Instagram несовместимыми с правами человека — например, удаление поста о политическом заключенном (Абдулле Оджалане) было сочтено ненужным и несоразмерным, поскольку его содержание представляло собой политическую пропаганду, не наносящую прямого вреда (). Такой анализ отражает то, что мог бы сделать суд, если бы статья 19 МПГПП была связана с Instagram.
- **Известные судебные дела:**Помимо иска Трампа, другие юридические или квази-юридические разбирательства рассматривали модерацию Instagram. В России, как уже отмечалось, суд назвал Meta «экстремистской» организацией отчасти из-за ее выбора модерации, фактически запретив Instagram по всей стране (Россия запретила Facebook и Instagram из-за закона об «экстремизме»). В ЕС новый Закон о цифровых услугах (DSA) требует от крупных платформ быть прозрачными в отношении своих алгоритмов и решений об удалении контента, а также предлагать пользователям апелляции — юридический толчок к надлежащей правовой процедуре в модерации. Хотя законы, такие как DSA, и различные

предлагаемые правила не касаются напрямую свободы слова, они направлены на проверку произвольной власти платформ, таких как Instagram, путем обязательного аудита их систем (который может выявить любые политические предубеждения или цензурные директивы в их алгоритмах). Мы также видим иски от отдельных лиц: например, некоторые консервативные влиятельные лица подали в суд за клевету или дискриминацию, когда их контент был помечен как дезинформация (утверждая, что эти ярлыки нанесли ущерб их репутации или бизнесу). Эти дела сталкиваются с трудными битвами, но указывают на тенденцию использования творческих юридических аргументов для оспаривания действий Instagram в отношении контента.

По сути, американское законодательство в настоящее время поддерживает право Instagram курировать контент, тогда как международные принципы прав человека призывают Instagram согласовывать свою политику с нормами свободы слова. Существует внутренний конфликт: Instagram может законно удалять речь, которая является законной (поскольку это частный форум), но делать это чрезмерно или неравномерно может нарушить дух прав человека. Как заметила Human Rights Watch в контексте удаления контента палестинцами, непоследовательное исполнение Meta привело к «ошибочному удалению контента» и непропорциональному подавлению одной точки зрения (Нарушенные обещания Меты: системная цензура палестинского контента в Instagram и Facebook | HRW), ущемляя права пользователей. Для решения этих проблем эксперты рекомендуют Instagram принять более четкие правила, обеспечить пропорциональность в применении и улучшить прозрачность и апелляции () (). Некоторые из этих шагов постепенно реализуются под давлением общественности и регулирующих органов.

Скрытые алгоритмы, теневые запреты и проблемы прозрачности

Возможно, самая коварная форма цензуры, предполагаемая в Instagram, — это не явное удаление контента, а скрытое алгоритмическое его ограничение. Алгоритм ленты Instagram и системы ранжирования контента могут эффективно сделать публикацию невидимой без ведома пользователя — явление, широко известное как теневой бан. Отсутствие прозрачности вокруг этих алгоритмов привело к широко распространенной обеспокоенности «невидимой» цензурой по политическим или даже коррупционным мотивам (например, предпочтение контента, который соответствует деловым интересам Instagram или внешнему давлению).

Понимание алгоритмического черного ящика: Instagram использует алгоритмы для отбора того, что видят пользователи — как в своей основной ленте, так и на вкладке «Обзор» или в результатах поиска по хэштегам. Эти алгоритмы настроены на максимизацию вовлеченности и безопасности пользователей, что означает, что они не просто показывают все от всех, на кого подписан пользователь, в хронологическом порядке. По замыслу они понижают рейтинг контента, на который, скорее всего, будет сообщено или который система подозревает в проблематичности (даже если он не был удален сразу). Например, если пост помечен искусственным интеллектом Instagram как потенциальная дезинформация, его распространение может быть сокращено на 80%, даже если пост остается видимым в профиле автора. Это создает сценарии, когда контент пользователя технически разрешен в Instagram, но почти никому не доходит, что фактически является формой цензуры. Это «теневой», потому что пользователь часто не уведомляется — его пост не удаляется, поэтому он не получает предупреждения, однако воздействие (резкое сокращение аудитории) реально (Теневой бан реален: как социальные сети решают, кому заткнуть рот - The Washington Post) (Теневой бан реален: как социальные сети решают, кому заткнуть рот - The Washington Post).

Расследования, такие как серия статей The Markup «Автоматизированная цензура», пролили свет на эти практики. The Markup обнаружил, что Instagram использует различные метки (спам, травля и т. д.) способами, которые влияют на способность пользователя апеллировать и видимость его контента.

В ходе тестов комментарии, которые Instagram помечал как «спам», приводили к мгновенным теневым банам — комментатор мог видеть свое замечание, но никто другой — тогда как комментарии, помеченные как **«преследование», давали пользователю возможность обжаловать удаление (Как мы исследовали теневой бан в Instagram – The Markup) (Как мы исследовали теневой бан в Instagram – The Markup). Это предполагает алгоритмическую стратегию: классифицируя определенные политические комментарии как не подлежащие обжалованию категории, например спам, платформа может тихо удалять их и предотвращать ответные действия. Действительно, во время удаления контента по израильско-палестинским вопросам в 2023 году многие пропалестинские комментарии были загадочным образом удалены как «спам», что не позволило пользователям оспорить это решение (Как мы исследовали теневой бан в Instagram – The Markup) (Как мы исследовали теневой бан в Instagram – The Markup). Такое непрозрачное использование алгоритмов можно рассматривать как преднамеренный (хотя и косвенный) метод цензурирования спорных точек зрения, сводя к минимуму негативную реакцию.

(Пониженный, удаленный и отклоненный: в Instagram есть не только теневой бан – The Markup) Рисунок: Иллюстрация, представляющая концепцию «теневого бана» в социальных сетях. Такие платформы, как Instagram, могут сделать пользователя практически невидимым, понижая рейтинг или скрывая его посты алгоритмически. (Теневой бан реален: как социальные сети решают, кому заткнуть рот - The Washington Post) (Теневой бан реален: как социальные сети решают, кому заткнуть рот - The Washington Post)

Политические или коррупционные мотивы:Продолжаются дебаты о том, является ли скрытая цензура Instagram политически мотивированной или обусловлена другими факторами (например, давлением рекламодателей или жалобами пользователей). Некоторые критики предполагают, что материнская компания Instagram, Meta, может настраивать алгоритмы, чтобы умилостивить могущественные правительства или группы интересов. Например, как отмечалось ранее, примирительные шаги Марка Цукерберга перед кампанией Трампа 2024 года (восстановление функций, которые повышают политический контент) последовали за явными угрозами и давлением со стороны республиканцев (Пользователи Instagram сталкиваются с заблокированными политическими хэштегами в результатах поиска) (Пользователи Instagram сталкиваются с заблокированными политическими хэштегами в результатах поиска). Это поднимает вопрос: были ли алгоритмы ранее настроены так, чтобы уменьшить видимость правого политического контента, а затем скорректированы, когда политические ветры изменились? Хотя Meta отрицает партийные манипуляции, сроки политических изменений могут показаться подозрительными. С другой стороны, защитники прав человека обвиняют Instagram в том, что он поддается правительственным запросам, граничащим с цензурой, — например, временно скрывает хэштеги, критикующие определенные режимы или события, по приказу властей. (Один пример за пределами нашей основной области: сообщения о том, что Instagram ограничил охват #EndSARS во время протестов в Нигерии после жалоб правительства, хотя компания сослалась на техническую ошибку.) Коррупция также может повлечь за собой преференциальное отношение: документы осведомителей (файлы Facebook) выявили внутреннюю программу под названием «XCheck», которая освобождала миллионы VIP-пользователей от стандартной модерации. Хотя XCheck был предназначен для предотвращения ложных срабатываний в важных аккаунтах, по сути, это означало, что некоторые элиты могли распространять контент (включая дезинформацию), за который обычные пользователи будут наказаны (). Такое неравное применение правил подрывает доверие — оно предполагает, что алгоритм может быть избирательным и беспринципным, более суровым к одним пользователям, более снисходительным к другим по причинам, не связанным с самим контентом.

Призывы к прозрачности и реформам:Среди экспертов формируется консенсус, что солнечный свет — лучшее средство от беспокойства о скрытой цензуре. Пропагандируются отчеты о прозрачности, независимые аудиты и инструменты расширения прав и возможностей пользователей. Пользователи хотят знать, как выразилась The Washington Post, «когда и почему [платформы] подавляют наши мегафоны» (Теневой бан реален: как социальные сети решают, кому заткнуть рот - The Washington Post). В ответ Instagram предпринял некоторые шаги: теперь он периодически публикует отчеты «Широко просматриваемый контент» и имеет функцию статуса аккаунта, которая

сообщает пользователям, рекомендуются ли их посты или нет. Но это только первые шаги. Закон о цифровых услугах в Европе заставит Instagram разрешить проверенным исследователям изучать его алгоритмы на предмет системных предубеждений. Кроме того, Совет по надзору Instagram призвал к более четким определениям — например, поручив Meta определить, что представляет собой «пограничный» контент, рейтинг которого понижается, и уведомлять пользователей, когда их охват сокращается.

Без таких мер остается подозрение, что скрытые алгоритмы могут быть использованы в качестве оружия для цензуры. Будь то сознательная настройка инженера или просто новое поведение систем модерации ИИ, эффект заключается в том, что определенные взгляды (часто те, которые бросают вызов статус-кво или «официальной» позиции) с трудом находят аудиторию в Instagram. В авторитарных странах эта динамика вызывает еще большее беспокойство: если Meta может ограничить контент в угоду правительству (или должен сделать это, чтобы избежать полного запрета), то права пользователей на свободу слова находятся во власти закулисных алгоритмических корректировок. В дальнейшем толчок к алгоритмической ответственности имеет решающее значение. Как показал один опрос, почти 1 из 10 пользователей социальных сетей в США считает, что они были подвергнуты теневому бану — реальному или нет, такое восприятие подрывает доверие (Проливая свет на теневой бан). Только посредством большей прозрачности, последовательного применения политики и возможностей для возмещения ущерба Instagram может развеять представление о том, что он управляет тайной цензурной машиной.

Заключение

Вышеприведенное исследование показывает, что подход Instagram к модерации контента неоднократно приводил к конфликту с принципами свободы слова и приводил к обвинениям в цензуре и дискриминации пользователей, придерживающихся «альтернативных» взглядов. От консервативных комментаторов и религиозных групп до активистов в зонах конфликта, широкий спектр пользователей столкнулся с удалениями, запретами или теневыми запретами, которые они воспринимали как несправедливые. В некоторых случаях Instagram действительно переусердствовал — как было показано, когда он ошибочно удалил безвредный контент (позже восстановив его) (Instagram удалил пост баптистского служения о том, почему пасторами должны быть только мужчины, а затем извинился | США), или когда расследования выявили систематическую предвзятость в правоприменении (например, несоразмерную цензуру пропалестинского контента) (Нарушенные обещания Меты: системная цензура палестинского контента в Instagram и Facebook | HRW). В других случаях платформа, возможно, применяла свои правила против разжигания ненависти или дезинформации законно, однако последствия ощущались по идеологическим линиям, подпитывая повествование о преднамеренном замалчивании. Пандемия COVID-19 и выборы в США 2020 года, в частности, продемонстрировали готовность Instagram встать на сторону того, что он считал правдой, а не ложью, — даже ценой отчуждения тех, кто придерживался противоположных мнений. Это подняло глубокие вопросы: должна ли частная компания с охватом Instagram иметь право решать, какие политические или научные идеи приемлемы? И если она действительно обладает такой властью, какие проверки должны существовать для предотвращения злоупотреблений или ошибок?

Юридически Instagram имеет право модерировать контент (защищенное такими законами, как Раздел 230 в США), но суд общественного мнения и международные стандарты прав человека предъявляют к нему более высокие требования. Свобода слова ценится во всем мире как одно из основных прав, и хотя Instagram не является правительством, его платформа стала де-факто публичной площадкой для дискурса. Когда пользователи блокируются или подвергаются теневому бану, их возможность участвовать в этой публичной площадке затрудняется, что вызывает обеспокоенность по поводу прав человека. Документы и свидетельства (например, письмо Цукерберга в Конгресс) даже предполагают, что порой цензура Instagram поощрялась правительственными чиновниками (Цукерберг заявил, что администрация Байдена оказала давление на Meta, чтобы тот подверг цензуре

контент о COVID-19 | Reuters) (Цукерберг заявил, что администрация Байдена оказала давление на Meta, чтобы тот подверг цензуре контент о COVID-19 | Reuters), стирая грань между частной модерацией контента и государственной цензурой через доверенных лиц.

Чтобы смягчить эти проблемы, обсуждаются различные решения. К ним относятся: более четкие Руководящие принципы сообщества, применяемые равномерно по всем идеологиям; улучшенные процессы апелляции и прозрачность, чтобы пользователи знали, почему были предприняты действия, и могли их оспорить (больше никаких необъяснимых ярлыков «спам» на политических выступлениях); регулярные проверки алгоритмических результатов для выявления непреднамеренной предвзятости; и, возможно, сторонний надзор для проверки спорных решений модерации (расширение модели Совета по надзору). Со стороны пользователей наблюдается стремление к повышению осведомленности о цифровых правах, напоминая людям, что, хотя свобода слова является правом, частные платформы также имеют права — отсюда важность создания альтернативных каналов и отказа от опоры на единственный корпоративный форум для выражения своего мнения.

Instagram находится на распутье, типичном для крупных технологических компаний: он должен сбалансировать безопасность, правду и корпоративные ценности с беспорядочной, иногда оскорбительной сферой свободы слова. История его споров о контенте — от жалоб консерваторов до международных инцидентов — показывает, что слишком большая ошибка в ту или иную сторону вызывает ответную реакцию. В ближайшие годы есть надежда, что благодаря большей прозрачности и соблюдению принципиальных политик, основанных на правах человека, Instagram сможет снизить восприятие дискриминации точек зрения. Обеспечение того, чтобы пользователи всех мастей чувствовали, что их голос может быть услышан (в рамках закона и уважения), имеет важное значение для легитимности платформы. Задача колоссальная: создать онлайн-среду, которая будет одновременно цивилизованной и открытой, где будет сдерживаться вредоносная дезинформация, но будут процветать здоровые дебаты — включая несогласные, альтернативные взгляды. Сможет ли Instagram ответить на этот вызов, существенно повлияет на глобальный дискурс в социальных сетях и свободу слова в цифровую эпоху.

Источники:

1. Christian Post – Instagram приносит извинения за удаление поста баптистского служения (2018) (Instagram удалил пост баптистского служения о том, почему пасторами должны быть только мужчины, а затем извинился | США) (Instagram удалил пост баптистского служения о том, почему пасторами должны быть только мужчины, а затем извинился | США)
2. Европейский консерватор – Instagram блокирует аккаунты французских правых (август 2024 г.) (Instagram закрывает аккаунты нескольких французских правых голосов—Европейский Консерватор) (Instagram закрывает аккаунты нескольких французских правых голосов—Европейский Консерватор)
3. Reuters – Письмо Цукербергу: Белый дом оказал давление на Meta, чтобы тот подверг цензуре контент о COVID (Цукерберг заявил, что администрация Байдена оказала давление на Meta, чтобы тот подверг цензуре контент о COVID-19 | Reuters) (Цукерберг заявил, что администрация Байдена оказала давление на Meta, чтобы тот подверг цензуре контент о COVID-19 | Reuters)
4. Reuters – Meta допускает призывы к насилию против российских оккупантов (март 2022 г.) (Facebook разрешает публиковать военные посты, призывающие к насилию против российских оккупантов | Reuters) (Facebook разрешает публиковать военные посты, призывающие к насилию против российских оккупантов | Reuters)
5. NPR – Meta урегулировала вопрос с Трампом из-за запрета Facebook/Instagram, заплатив 25 миллионов долларов (январь 2025 г.) (Meta согласилась выплатить Трампу 25 миллионов долларов для урегулирования иска из-за блокировок Facebook и Instagram: NPR) (Meta согласилась выплатить Трампу 25 миллионов долларов для урегулирования иска из-за блокировок Facebook и Instagram: NPR)

6. Reuters – Трамп подает в суд на социальные сети, обвиняя их в предвзятости, оспаривая раздел 230 (июль 2021 г.) (Трамп подал в суд на Facebook, Twitter и Google, заявив о цензуре | Reuters) (Трамп подал в суд на Facebook, Twitter и Google, заявив о цензуре | Reuters)
7. Axios – Instagram заблокировал политические хэштеги, пользователи заявляют о предвзятости (январь 2025 г.) (Пользователи Instagram сталкиваются с заблокированными политическими хэштегами в результатах поиска) (Пользователи Instagram сталкиваются с заблокированными политическими хэштегами в результатах поиска)
8. Washington Post – Теневой бан в социальных сетях и призывы к прозрачности (Теневой бан реален: как социальные сети решают, кому заткнуть рот - The Washington Post) (Теневой бан реален: как социальные сети решают, кому заткнуть рот - The Washington Post)
9. *Разметка*– Расследование показало, что Instagram понизил рейтинг военного контента и скрыл комментарии (февраль 2024 г.) (Пониженный, удаленный и отклоненный: в Instagram есть не только теневой бан – The Markup)
10. NPR – Instagram забанил Роберта Ф. Кеннеди-младшего за дезинформацию о вакцинах (Instagram заблокировал РФК-младшего за распространение дезинформации о вакцинах : Обновления о коронавирусе : NPR) (Instagram заблокировал РФК-младшего за распространение дезинформации о вакцинах : Обновления о коронавирусе : NPR)
11. Human Rights Watch – цензура пропалестинского контента со стороны Meta (декабрь 2023 г.) (Нарушенные обещания Меты: системная цензура палестинского контента в Instagram и Facebook | HRW)





